

中国宏观经济内生增长因素分析^{*}

——基于地理加权回归 (GWR) 模型的实证分析

韦米佳

内容提要: 本文采用 2006 年中国大陆 30 个省、市、自治区的宏观经济数据, 应用空间计量的莫兰指数 (Moran's I)^① 以及地理加权回归模型, 对中国宏观经济增长的决定因素进行分析。实证结果表明, 在中国大陆资本对经济增长的贡献大于劳动对经济增长的贡献, 而劳动力受教育水平对地区经济的影响则非常小。

关键词: 空间效应 内生经济增长 地理加权回归模型

作者: 厦门大学经济学院财政系博士研究生。

一、引言

二十世纪八十年代末兴起的内生经济增长理论通过将知识积累过程引入典型的增长模型, 一举突破了长期困扰新古典增长理论的技术外生性问题, 实现了技术进步的内生化, 这意味着技术进步不再受经济系统之外变量的控制, 而是由 R&D 和教育等经济系统内部因素所决定。从目前国内外的研究成果来看, 几乎所有有关经济增长机制的实证研究都假设截面单元是同质的, 即不同的地区之间没有差异, 事实上, 影响经济增长的要素在不同区域是有区别的, 不同区域间的投入要素做简单的“同质化”处理值得商榷。同时, 相关研究的计量方法也主要是传统的回归分析方法 (如多元统计分析、回归分析、数据包络分析等), 未考虑区域之间的空间关联, 而空间计量经济学认为区域之间的经济行为会相互影响, 导致地区之间的经济行为存在溢出效应, 例如经济产出不仅受本地投资强度的影响, 还会受到周边地区投资活动产生的溢出效应及政策的影响。吴玉鸣、李建霞 (2006) 利用地理加权回归模型就发现省域工业全要素生产率具有空间效应。

本文选用教育来替代技术内生性。教育投资可以促进人力资本积累和技术进步, 在经济增长中的作用是不可替代的, 对经济增长不仅有“水平效应”, 而且有“增长效应”。为解决经济增长的空间效应以及各省份不同质两大问题, 本文采用考虑了空间效应的地理加权回归模型对经济增长问题进行分析。

^{*} 本文获厦门大学经济与管理教学实验中心开放性实验项目资助, 在此表示感谢。

^① 莫兰指数 (Moran's I) 是由帕特里克·莫兰 (Patrick A. P. Moran) 提出的, 用来衡量空间自相关的一种方法。

二、模型的理论依据

新古典主义经济学认为如果一个地区的资本投入、劳动力投入越多,而且该地区的技术水平越高,则该地区的 GDP 就会越高。影响经济增长的因素很多,这里仅考虑资本投入、劳动力投入以及劳动力所受的教育水平,则假设如下:

假设 1: 地方资本投资越多,其 GDP 越高;

假设 2: 地方劳动力数量越多,其 GDP 越高;

假设 3: 地方劳动力受教育水平越高,其 GDP 越高。

采用纳入人力资本的索洛模型,同时综合内生经济增长的干中学理论^①,设定生产函数为柯布—道格拉斯 (Cobb-Douglas) 生产函数,其数学表达式如下:

$$Y(t) = K(t)^\alpha [A(t)H(t)]^\beta, A(t) = BK(t)^\phi, H(t) = L(t)G(E), G(E) = e^{\varphi E}, \varphi > 0 \quad (1)$$

其中 Y 代表经济产出; K 代表资本投入; A 代表技术; H 代表包含人力资本的劳动力; G 代表劳动力包含的人力资本; E 代表工人接受教育的平均年限; e 为自然对数。将技术函数和包含人力资本的劳动力函数代入生产函数,可得:

$$Y(t) = B^\beta K(t)^{\alpha + \beta\phi} L(t)^\beta e^{\beta\varphi E} \quad (2)$$

从上式可以看出,经济产出取决于资本、劳动力数量以及劳动力接受教育的时间,为了便于估计,对 (2) 两边取对数,得到:

$$\ln Y(t) = \beta \ln B + (\alpha + \beta\phi) \ln K(t) + \beta \ln L(t) + \beta\varphi E \quad (3)$$

三、GWR 模型

当建立模型使用的是横截面数据时,如果经济数据在空间上表现出的依赖性和异质性使得解释变量对被解释变量的影响在不同区域之间呈现差异性,那么使用 GWR 模型是合理的选择。

在 GWR 模型中,特定区位的回归系数不再是利用全部信息获得的假定常数,而是利用邻近观测值的子样本数据信息进行局域回归估计而得的、随空间地理位置变化而变化的变数, GWR 模型可以表示为:

$$y_i = \beta_0(u_i, v_i) + \sum_{j=1}^k \beta_j(u_i, v_i) x_{ij} + \epsilon_i \quad (4)$$

β_j 表示与观测值联系的待估计参数向量,是关于地理位置 (u_i, v_i) 的 $k+1$ 元函数; ϵ 是第 i 个区域的随机误差,满足零均值、同方差、相互独立等假定。模型中,每个区域都有一个对应的估计函数,其对数似然函数可以表示为:

$$\begin{aligned} \log L &= L[\beta_0(u, v), \dots, \beta_k(u, v) | M] \\ &= -\frac{1}{2\sigma^2} \sum_{i=1}^n [y_i - \beta_0(u_i, v_i) - \sum_{j=1}^k \beta_j(u_i, v_i) x_{ij}]^2 + \alpha \end{aligned} \quad (5)$$

式中, α 为常数, $M = [y_i, x_{ij}, (u_i, v_i), i = 1, 2, \dots, n, j = 1, 2, \dots, k]$ 。由于极大似然法的解不是唯一的,黑斯蒂和蒂鲍斯兰尼 (Hastie and Tibshirani, 1993) 认为用该方法求解不恰

^① 该理论的核心思想是:个人在生产过程中会考虑生产过程的改进方法,因此技术创新是传统经济活动的副产品;所有的资源都用于生产,学习是生产新资本的副产品,知识存量是资本存量的函数;劳动力本身包含着人力资本,取决于劳动者所受教育的年限,并且人力资本具有竞争性和排他性。

当,提出了局域求解法,原理与方法如下:

对于第 s 个空间位置 $[(u_s, v_s), s = 1, 2, \dots, n]$, 任取一空间位置 (u_0, v_0) 与其位置邻近, 构造一个简单的回归模型:

$$y_i = \gamma_0 + \sum_{j=1}^k \gamma_j x_{ij} + \epsilon_i \quad (6)$$

γ_j 为 GWR 模型中 $\beta_j(u_s, v_s)$ 的近似值, 是一常数, 通过考虑与点 (u_0, v_0) 相邻近的点来校正经典回归模型中的解, 常用的方法是加权最小二乘法, 寻找合适的 γ_j 使得下式最小:

$$\sum_{i=1}^k W(d_{0i})(y_i - \gamma_0 - \sum_{j=1}^k \gamma_j x_{ij})^2 \quad (7)$$

式中, d_{0i} 为位置 (u_0, v_0) 和 (u_i, v_i) 的空间距离, $W(d_{0i})$ 为空间权值。令 $\hat{\gamma}_j$ 为 $\hat{\beta}_j(u_s, v_s)$ 的估计值, 可得 GWR 模型在空间位置 (u_s, v_s) 上的估计值 $\{\hat{\beta}_0(u_s, v_s), \hat{\beta}_1(u_s, v_s), \dots, \hat{\beta}_k(u_s, v_s)\}$ 。对上式求 γ_j 的一阶偏导数, 并令其等于 0, 可得:

$$\hat{\gamma}_j = (X'W_0^2X)^{-1}(X'W_0^2Y) \quad (8)$$

W_0 为 $[W(d_{01}), W(d_{02}), \dots, W(d_{0n})]$ 的对角线矩阵。可以看出 $\hat{\beta}_j(j = 1, 2, \dots, k)$ 的 GWR 估计值是随着空间权值矩阵 W_{ij} 的变化而变化的, 因此 W_{ij} 的选择至关重要, 一般由观测值的空间(经纬度)坐标决定。

实际研究中常用的空间距离权值计算公式有三种 (LeSage, 2004)。

(1) 高斯距离权值 (Gaussian Distance)

$$W_{ij} = \Phi(d_{ij}/\sigma\theta)$$

(2) 指数距离权值 (Exponential Distance)

$$W_{ij} = \sqrt{\exp(-d_{ij}/q)}$$

(3) 三次方距离权值 (Tricube Distance)

$$W_{ij} = [1 - (\theta/d_{ij})^3]^3$$

其中 d_{ij} 为第 i 个区域与第 j 个区域间的地理距离, Φ 为标准正态分布密度函数, q 为观测值 i 到第 q 个最近区域之间的距离, σ 为距离向量的标准差, θ 为衰减参数 (窗宽)。

在空间权值矩阵中, d 和 θ 非常关键。如果 d 较大, 则局域模型的解越趋向于全域模型的解; 如果 d 等于所研究空间任意两点间的最大距离, 则全域和局域两个模型将相等, 反之则相反。若 θ 趋于无穷大, 任意两点的权重将趋于 1, 则 GWR 就等于以 OLS 估计的经典线性回归; 反之, 参数估计将更加依赖于邻近的观测值。

四、实证分析

(一) 数据搜集。

本文选取中国大陆 30 个省、市、自治区^① 2006 年的数据为样本进行实证分析, 数据来源于《中国统计年鉴 2007》和《中国劳动统计年鉴 2007》。这里选取的被解释变量为地区生

^① 西藏由于数据统计方面的原因排除在外。

产总值，解释变量包括各地区资本形成总额、就业人员总数和就业人员平均受教育年限^①。

(二) 空间相关性检验与分析。

在建立模型分析研究之前，先进行空间相关性预检验。如果空间效应在发挥作用，则需要将空间效应纳入模型分析框架之中，并采用适合于空间计量的方法进行估计；反之，则可采用一般估计方法估计模型参数。在实际的空间相关应用研究中，莫兰指数是常用方法，本文也采用该方法检验区域经济发展变量的空间相关性存在与否。

莫兰指数计算公式如下：

$$Moran's I = \frac{\sum_{i=1}^n \sum_{j=1}^n W_{ij} (Y_i - \bar{Y})(Y_j - \bar{Y})}{S^2 \sum_{i=1}^n \sum_{j=1}^n W_{ij}}$$

(9)

其中， $S^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2$ ， $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$ ， Y_i 表示第 i 个地区的观测值， W 为空间权值矩阵，一般采用邻近标准，其目的是定义空间对象的相互邻近关系，两个地区相邻取值为 1，否则为 0。利用公式(5)检测 2006 年中国大陆 30 个省份经济增长在地理空间上的相关性。

经济增长的莫兰指数为 0.58322696，正态统计量为 2.77379194^②，表明中国 30 个省、市和自治区的经济增长值在空间分布上具有明显的正相关性，正的空间相关性代表相邻地区的空间联系结构特性类

表 1 莫兰指数 检测结果

变量	Moran's I	Moran's I-statistic	P 值	显著程度
GDP	0.58322696	2.77379194	(0.001)	***

似，即地区生产总值较高的省区相互靠近，地区生产总值较低的省区相互靠近。因此，从整体上讲省区之间的经济生产总值是存在空间相关性的，也就是说存在着空间上明显的集群现象，所以有必要在估计时采用纳入空间依赖性的空间计量经济模型进行估算。

(三) 回归估计和分析。

实证分析选取的工具为 Matlab7.0。首先，对 2006 年中国大陆 30 个省份的地区生产总值、地区资本形成总额、就业人员总数和就业人员人均受教育年限等经济数据进行最小二乘回归估计（OLS 估计），估计结果见图 1 和表 2：

表 2 OLS 估计结果

Ordinary Least-squares Estimates Dependent Variable = GDP R-squared = 0.9853 Rbar-squared = 0.9836 sigma^2 = 0.0135 Durbin-Watson = 2.0163 Nobs, Nvars = 30, 4	Variable	Coefficient	t-statistic	t-probability
	constant	0.084105	0.371350	0.713386
	lnK	0.568712	5.530069	0.000008
	lnL	0.491814	5.240079	0.000018
	lnG(E)	0.072544	3.145847	0.004119

① 就业人员平均受教育年限由各地区不同类别就业人员受教育年限加权而得，计算公式为： $G(E) = \sum_{i=1}^n E_i P_i$ ，其中 E_i 为第 i 种就业人员受教育年限， p_i 为第 i 种就业人员占总就业人员比例。

② 大于正态分布函数在 0.01 水平下的临界值 1.96。

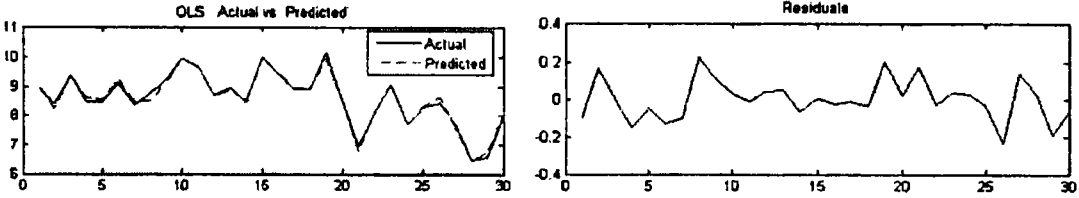


图 1 OLS 估计结果

由表 2 可知,生产函数的拟合优度达到 0.9853,模型整体上通过了 5 % 水平的显著性检验,但常数项未能通过。之前的莫兰指数检验已证实中国大陆 30 个省、市、自治区的经济发展具有明显的空间效应,这些都说明经典 OLS 估计存在问题。经分析,问题产生于 OLS 得到的回归系数是全局的,各区域整体上被假定是同一个常数,无法区分各要素在局部省份对经济增长的影响。为了解决这个问题,采用地理加权回归模型(GWR)进行回归,得到的结果如下:

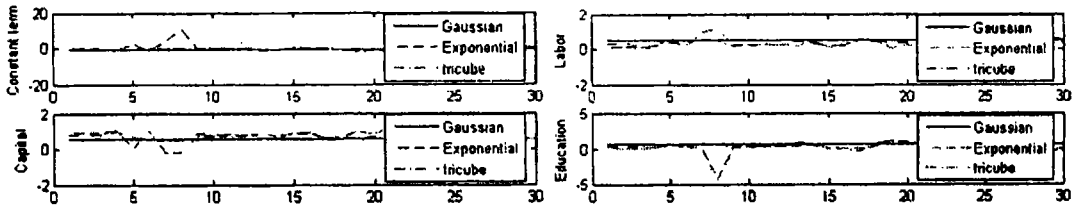


图 2 GWR 各参数的估计值

图 2 分别给出了高斯距离权值(Gaussian Distance)、指数距离权值(Exponential Distance)、三次方距离权值(Tricube Distance)的地理加权回归模型各参数的估计值,各回归系数均通过了 5%水平下的变量显著性检验,这表明考虑了空间效应的地理加权回归模型整体拟合效果更好。

表 3 回归结果拟合优度的比较

回归方法	.OLS	GWR (Gaussian Distance)	GWR (Exponential Distance)	GWR (Tricube Distance)
R ²	0.9853	0.9856	0.9992	0.9923

表 4 GWR 估计结果 (指数距离权值, Exponential Distance)

地区	constant	lnK	lnL	lnG(E)	地区	constant	lnK	lnL	lnG(E)
北 京	0.279649	0.899926	0.15039	0.0163	河 南	0.066972	0.813973	0.269310	0.022545
天 津	0.426096	0.898623	0.139449	0.010781	湖 北	0.099177	0.619756	0.520044	0.016040
河 北	0.35509	0.921149	0.124462	0.009722	湖 南	0.343996	0.686803	0.429525	0.063783
山 西	0.74837	1.058007	0.087865	0.029658	广 东	0.881311	0.971100	0.086236	0.156550
内 蒙 古	3.62391	0.003078	0.682798	0.048959	广 西	0.342910	0.797103	0.246424	0.121839
辽 宁	0.766465	1.040323	0.090536	0.042207	海 南	0.698754	1.213675	0.149207	0.102772

地区	constant	lnK	lnL	lnG(E)	地区	constant	lnK	lnL	lnG(E)
吉 林	3.539765	0.203788	0.967883	0.017587	重 庆	0.008089	0.617255	0.522269	0.005723
黑龙江	6.557020	0.152290	1.149457	0.502877	四 川	0.484671	0.458128	0.608178	0.027283
上 海	0.52023	0.746833	0.259687	0.051011	贵 州	0.082633	0.533229	0.520178	0.082802
江 苏	0.576989	0.760834	0.243952	0.043754	云 南	0.188885	0.360571	0.711754	0.123992
浙 江	0.410972	0.769954	0.248052	0.050607	陕 西	0.155044	0.782715	0.343517	0.007053
安 徽	0.396917	0.733589	0.291837	0.048797	甘 肃	1.198282	0.724334	0.392563	0.149791
福 建	0.151087	0.880947	0.208638	0.042310	青 海	1.437458	0.838662	0.266077	0.187400
江 西	0.540773	0.734390	0.398189	0.067487	宁 夏	0.932319	0.369150	0.728925	0.054021
山 东	0.618879	0.885306	0.134349	0.007094	新 疆	0.046406	0.351136	0.287946	0.404813

以拟合优度最好的以指数距离权值进行估计的 GWR 回归结果为例（如表 4），我们可以发现，无论是资本投入、劳动力投入，还是劳动力受教育水平的回归系数都显著为正，这表明文章前面所作的 3 个假设都是正确的，即地方资本投资越多，劳动力数量越多，而且劳动力受教育水平越高，其 GDP 越高。

但是在各个不同的省份，各种投入要素的贡献是不一样的。在中国大陆 30 个样本省份中，只有内蒙古、吉林、黑龙江、四川、云南和宁夏 6 个省份劳动对经济增长的贡献大于资本对经济增长的贡献，其余 24 个省份都是资本对经济增长的贡献大于劳动对经济增长的贡献，这表明我国目前经济增长主要还是靠增加资本投入，包括增加国内固定资产投资和引进外资等，而不是我们通常认为的主要是丰富的劳动力资源带动了我国经济的发展。在回归结果中我们还可以发现，劳动力受教育水平对地区 GDP 的影响普遍来说都比较小，且远小于资本的影响或劳动力数量本身的影响，这与中国高科技产业不发达，产业组织结构比较落后的现状比较吻合。

五、总 结

实证分析发现，在中国大陆资本对经济增长的贡献大于劳动对经济增长的贡献，而劳动力受教育水平对地区经济的影响远小于资本或者劳动力数量的影响。随着金融危机的爆发，引进外资大幅下降，劳动密集型和外向型经济受到了严重的影响，中国经济转变增长方式，进行产业结构升级刻不容缓，而进行自主创新，培育优质的人力资源，发展高新技术产业是关键，政府应想方设法促使教育改革成果真正带动中国经济发展。

实证结果表明，考虑了空间效应的地理加权回归模型在拟合优度和参数检验方面都好于传统的最小二乘回归模型，这表明省际宏观经济发展确实存在着地理空间上的关联，具有显著的空间效应。GWR 模型得到的回归系数有着显著的差异，表明影响各省、市、自治区经济增长的因素具有较明显的区别，实证结果也更为可信。

地理加权回归模型是处理具有空间效应的区域经济发展问题的一个好方法，本文仅考虑了各区域实际地理位置上的关联。在现代信息社会里，两个地区实际联系已经不再局限于地

理上的相连，贸易协议的签订、政策上的相关性等会形成虚拟空间上的关联，这方面的研究有待专家学者进一步深入。

参考文献：

1. Anselin L.Spatial econometrics(1988):methods and models[M].Dordrecht ,Kluwer Academic Publishers.
2. Lesage J P,Luc Anselin,Raymond J G,Florax,Sergio J Rey(2004).A family of geographically weighted regression models in advances in spatial econometrics [M].Berlin:Springer-Verlag.
3. 吴玉鸣、李建霞：《基于地理加权回归模型的省域工业全要素生产率分析》，《经济地理》2006 年第 5 期。
4. 顾佳峰：《中国教育支出与经济增长的空间实证分析》，《教育与经济》2007 年第 1 期。

作者单位：厦门大学经济学院财政系
邮 编：361005

书 讯
祝 健著：

《中国农村金融体系重构研究》
社会科学文献出版社 2008 年出版

作者围绕中国农村金融体系演变、改革、发展的主线，从金融基础、金融支持、金融效率等方面分析农村金融供给抑制现状及其原因，揭示“三农”发展金融体系支持的内在逻辑与现实依据，确立重构农村金融体系的目标与原则。在此基础上，探讨商业性金融、政策性金融和合作性金融的运行原理及支农效应，提出深化农村金融机构改革、培育竞争性金融市场、创新金融服务产品、强化金融制度供给与金融监管的相关构想。全书 28 万字。文字流畅，创意颇多，应用性强。本书是福建师范大学原校长、福建师范大学马克思主义研究院院长李建平教授主编的大型系列丛书“马克思主义理论与现实研究文库”（计划 10 年左右出十辑 100 本书）中之一本，属第二辑《〈资本论〉与马克思主义经济理论研究专集》。此辑已出版的还有黄茂兴等的《技术选择与产业结构升级》、《改革开放 30 年中国经济热点的回眸与展望》、杨强的《中国个人收入的公平分配》等著作。文库有李建平教授写的总序，昭示本文库是福师大学习、研究、宣传马克思主义理论的重要阵地，也是开展对外学习交流的重要平台，所追求的终极目的是“为马克思主义在中国的发展，为全面建设小康社会、开创中国特色社会主义新局面作出新的更大的贡献。”（张文）